

Parabolic patterns in the scatter plot of Pythagorean triples

Kuan Peng titang78@gmail.com

15 August 2025

Abstract: The scatter plot of Pythagorean triples exhibits distinct parabolic patterns whose origins have not been fully characterized. In this work, we derive explicit parabolic functions directly from the Pythagorean equation and demonstrate their correspondence with these patterns. The analysis shows that basic Pythagorean triples are regularly distributed on the (X,Y) plane, occurring precisely at the intersections of horizontal and vertical parabolas. The derived functions align closely with the observed parabolic structures, and a density analysis of the parabolas explains the prominence of these patterns in the scatter plot of all Pythagorean triples.

1.	INTRODUCTION.....	1
2.	FUNCTIONS FOR THE PARABOLIC PATTERNS.....	2
a)	<i>Derivation from Pythagorean equation.....</i>	2
b)	<i>Focal point of the parabolas</i>	3
3.	SCATTER PLOT OF BASIC PYTHAGOREAN TRIPLES	3
a)	<i>Functions for basic Pythagorean triples</i>	4
b)	<i>Scatter plot of basic Pythagorean triples</i>	5
c)	<i>Comparison with the table of basic Pythagorean triples</i>	5
4.	SCATTER PLOT OF ALL PYTHAGOREAN TRIPLES	6
a)	<i>Multiple Pythagorean triples.....</i>	6
b)	<i>Scatter plot.....</i>	6
c)	<i>Video of the plotting</i>	7
5.	WHY THE PARABOLIC PATTERNS SHOW UP.....	7
6.	DISCUSSION.....	8
	LETTER TO READERS	9

1. Introduction

Pythagorean triples, such as $(3, 4, 5)$, $(5, 12, 13)$, $(8, 15, 17)$, $(7, 24, 25)$, $(20, 21, 29)$, $(12, 35, 37)$, $(9, 40, 41)$, $(28, 45, 53)$, $(11, 60, 61)$, $(33, 56, 65)$, $(48, 55, 73)$, $(39, 80, 89)$, $(65, 72, 97)$, have long been recognized for their mathematical elegance. At first glance, these triples appear unordered, and until recently, no method existed to arrange them in a systematic and visually coherent manner.

In a previous work, « [Classification of Pythagorean triples](#) and [reflection on Fermat's last theorem](#)», we proposed a method for classifying all Pythagorean triples. In this approach, triples are organized into a table in which columns correspond to constant $Z-X$ values and rows correspond to constant $Z-Y$ values (Table 1). All of the triples listed above appear in this table.

k=1	Z-X = 2			Z-X = 8			Z-X = 18			Z-X = 32			Z-X = 50		
Z-Y = 1	3	4	5	5	12	13	7	24	25	9	40	41	11	60	61
Z-Y = 4	8	6	10	12	16	20	16	30	34	20	48	52	24	70	74
Z-Y = 9	15	8	17	21	20	29	27	36	45	33	56	65	39	80	89
Z-Y = 16	24	10	26	32	24	40	40	42	58	48	64	80	56	90	106
Z-Y = 25	35	12	37	45	28	53	55	48	73	65	72	97	75	100	125

Table 1

Another approach for representing Pythagorean triples is to plot them as points in a Cartesian coordinate system, with the abscissa and ordinate corresponding to the X and Y values of each triple. When many triples are plotted, the resulting collection of points forms a scatter plot, referred to here as the scatter plot of Pythagorean triples.

An example of such a plot is shown in the [Pythagorean triple](#) article on Wikipedia (Figure 1). This scatter plot exhibits striking parabolic patterns, suggesting an underlying geometric order to the triples. The “[Distribution of triples](#)” section of the article describes these patterns as follows:

“*Within the scatter, there are sets of parabolic patterns with a high density of points and all their foci at the origin, opening up in all four directions. Different parabolas intersect at the axes and appear to reflect off the axis with an incidence angle of 45 degrees, with a third parabola entering in a perpendicular fashion.*”

Although these patterns have been noted, no explanation for their origin has previously been established.

Building upon the theoretical framework developed in « [Classification of Pythagorean triples](#) and [reflection on Fermat's last theorem](#) », we derive four explicit parabolic functions from the Pythagorean equation that account for these patterns. The four parabolas plotted in Figure 2 represent these functions. At the intersection points along the Y-axis, the features described above are precisely observed: parabolas intersecting the axes, reflecting at 45 degrees, and intersected perpendicularly by a third parabola.

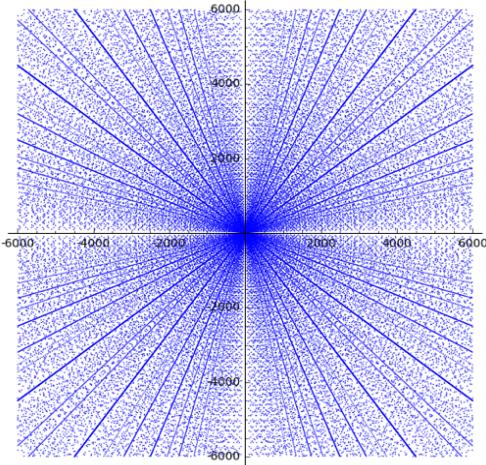


Figure 1

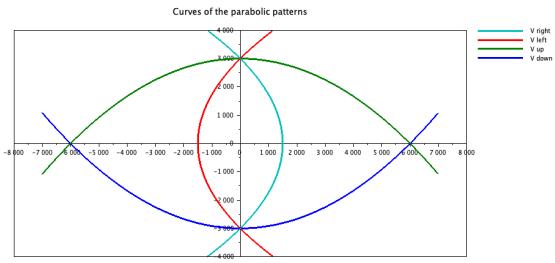


Figure 2

2. Functions for the parabolic patterns

a) Derivation from Pythagorean equation

These parabolic patterns are certainly a feature of Pythagorean equation which we write below:

$$X^2 + Y^2 = Z^2 \quad (1)$$

Because X and Y are squared, we can insert the sign \pm before X and Y and write Pythagorean equation as follow:

$$(\pm X)^2 + (\pm Y)^2 = Z^2 \quad (2)$$

For deriving the functions that describe the parabolic patterns we apply the change of variable proposed in « [Classification of Pythagorean triples](#) and [reflection on Fermat's last theorem](#) »:

$$Z = L \pm X \quad (3)$$

By introducing equation (3) into (2) we get :

$$\begin{aligned} Y^2 &= (L \pm X)^2 - X^2 \\ &= \pm 2XL + L^2 \end{aligned} \quad (4)$$

From equation (4) we derive the expression for X:

$$X = \pm \frac{Y^2 - L^2}{2L} \quad (5)$$

In the same way, we make the following change of variable for Y :

$$Z = K \pm Y \quad (6)$$

which we introduce into equation (2) and we get :

$$\begin{aligned} X^2 &= (K \pm Y)^2 - Y^2 \\ &= \pm 2YK + K^2 \end{aligned} \quad (7)$$

From equation (7) we derive the expression for Y :

$$Y = \pm \frac{X^2 - K^2}{2K} \quad (8)$$

Equation (8) is the two parabolic functions below :

$$Y = +\frac{X^2 - K^2}{2K} \quad (9)$$

$$Y = -\frac{X^2 - K^2}{2K} \quad (10)$$

Equation (5) is the two parabolic functions below :

$$X = +\frac{Y^2 - L^2}{2L} \quad (11)$$

$$X = -\frac{Y^2 - L^2}{2L} \quad (12)$$

Equation (9) represents a parabola that opens upward with its vertex at the bottom, while equation (10) defines a parabola that opens downward with its vertex at the top. Equation (11) corresponds to a parabola that opens to the right with its vertex on the left, and equation (12) defines a parabola that opens to the left with its vertex on the right. These four parabolas are plotted in Figure 2 and are labeled as V_{right} , V_{left} , V_{up} and V_{down} , where V denotes the vertex.

Equation (5) defines two parabolas whose axes of symmetry are horizontal; these are referred to as horizontal parabolas. Similarly, the parabolas defined by equation (8) have vertical axes of symmetry and are therefore called vertical parabolas.

b) Focal point of the parabolas

The equation that defines a parabola with its focal length f is given below, see [Parabola](#) :

$$y = \frac{x^2}{4f} \quad (13)$$

By comparing equation (8) with (13), we pose :

$$K = 2f \quad (14)$$

By introducing equation (14) into equation (8) we get :

$$Y = \pm \frac{X^2 - 4f^2}{4f} \quad (15)$$

The focal points of the parabolas described by this equation is on the Y-axis where X equals 0. We find the ordinate of the vertex of these parabolas by making X=0 in (15):

$$Y_v = \pm(0 - f) \quad (16)$$

The ordinate of the focal points of the parabolas equals :

$$\begin{aligned} Y_f &= Y_v \pm f \\ &= 0 \end{aligned} \quad (17)$$

So, the focal points of these two parabolas are at the origin :

$$\begin{aligned} X_f &= 0 \\ Y_f &= 0 \end{aligned} \quad (18)$$

In the same way, we find that the focal points of the parabolas defined by equation (5) are also at the origin. So, the focal points of the four parabolas are all situated at the origin. This has been noted by the description in the [“Distribution of triples”](#) which says: “all their foci at the origin.”

3. Scatter plot of basic Pythagorean triples

We distinguish two categories of Pythagorean triples: basic triples and multiple triples. Basic triples (X, Y, Z) satisfy $X^2 + Y^2 = Z^2$. So, the triple (kX, kY, kZ) is also a Pythagorean triple because :

$$(kX)^2 + (kY)^2 = (kZ)^2 \quad (19)$$

We call (kX, kY, kZ) multiple Pythagorean triples.

Basic triples are computed with equation (4) and (7). The parameters L and K are defined as below, see [« Classification of Pythagorean triples »](#) and [« reflection on Fermat’s last theorem »](#) :

$$L = 2b^2 \quad (20)$$

$$K = d^2 \quad (21)$$

with b and d being integers.

Then, the two vertical parabolic functions are:

$$Y = +\frac{X^2 - d^4}{2d^2} \quad (22)$$

$$Y = -\frac{X^2 - d^4}{2d^2} \quad (23)$$

The two horizontal parabolic functions are :

$$X = +\frac{Y^2 - 4b^4}{4b^2} \quad (24)$$

$$X = -\frac{Y^2 - 4b^4}{4b^2} \quad (25)$$

The Table 1 shows that the basic Pythagorean triples defined by the parameters L and those by K are equal.

a) Functions for basic Pythagorean triples

Basic Pythagorean triples defined with L are given by the following equations:

$$X = a^2 - b^2$$

$$Y = 2ab \quad (26)$$

The parameters a and b are defined as follow:

$$b=1, 2, 3, 4\dots \quad (27)$$

$$a = b+j \quad (28)$$

with j being integer.

One value of the parameter b corresponds one parabola; one j corresponds one basic Pythagorean triple. We compute one parabola with one value of b , then, the following parabola with $b+1$ and so on. Then, we compute all the basic Pythagorean triples on one parabola with j . The limit values for j vary with b and are computed with equation (26). Horizontal parabolas intersect the Y-axis at two points where X equals zero, so :

$$a^2 - b^2 = 0 \quad (29)$$

We introduce equation (28) into equation (29) and obtain:

$$\begin{aligned} a^2 - b^2 &= (a+b)(a-b) \\ &= (b+j+b)(b-j-b) \quad (30) \end{aligned}$$

$$\begin{aligned} &= j(2b+j) \\ j(2b+j) &= 0 \quad (31) \end{aligned}$$

This equation has two solutions for j :

$$j = 0, \quad j = -2b \quad (32)$$

We plot the basic Pythagorean triples for j within the following interval :

$$-2b \leq j \leq 0 \quad (33)$$

The basic Pythagorean triples defined with K are on vertical parabolas. Their X and Y are given by the following equations:

$$\begin{aligned} X &= cd \\ Y &= \frac{c^2 - d^2}{2} \quad (34) \end{aligned}$$

with :

$$d=1, 2, 3, 4\dots \quad (35)$$

$$c = d+2\cdot i \quad (36)$$

These parabolas intersect the X-axis at two points where Y equals zero, so:

$$c^2 - d^2 = 0 \quad (37)$$

We introduce equation (36) into equation (37) and obtain:

$$\begin{aligned} c^2 - d^2 &= (d + c)(d - c) \\ &= (d + d + 2i)(d - d - 2i) \quad (38) \\ &= -2i(2d + 2i) \end{aligned}$$

$$2i(2d + 2i) = 0 \quad (39)$$

Equation (39) has two solutions for i :

$$i = 0, \quad i = -d \quad (40)$$

We plot the basic Pythagorean triples for i within the following interval :

$$-d \leq i \leq 0 \quad (41)$$

b) Scatter plot of basic Pythagorean triples

The basic Pythagorean triples on vertical parabolas are computed with equation (26) with the parameters below :

$$b=1, 2, 3, 4 \dots \text{ and } -2b \leq j \leq 0 \quad (42)$$

The basic Pythagorean triples on horizontal parabolas are computed with equation (34) with the parameters below :

$$d=1, 2, 3, 4 \dots \text{ and } -d \leq i \leq 0 \quad (43)$$

We plot these basic Pythagorean triples and their parabolas in Figure 3.

We notice that in the first quadrant the blue dots cover exactly the red dots, confirming that the basic Pythagorean triples on the vertical and horizontal parabolas are indeed the same Pythagorean triples. We notice also that basic Pythagorean triples are at the intersections of the horizontal and vertical parabolas.

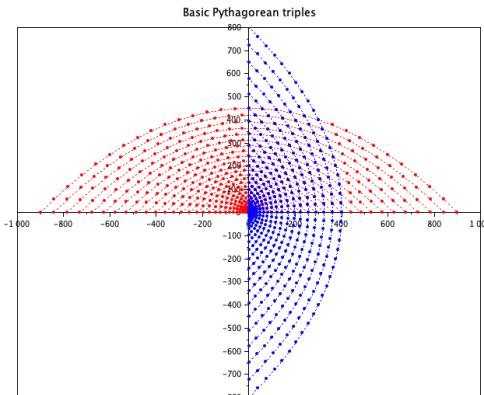


Figure 3

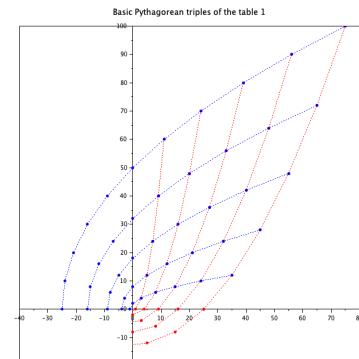


Figure 4

c) Comparison with the table of basic Pythagorean triples

We have plotted in Figure 4 the scatter plot of the basic Pythagorean triples shown in Table 1 with the parabolas they belong to. The blue parabolas are computed with equation (12) for $b=1, 2, 3, 4, 5$ and $j=-5, -4 \dots 4, 5$. The red parabolas are computed with equation (10) for $d=1, 2, 3, 4, 5$ and $0 \leq i \leq \text{int}(d/2)$. The dots in the first quadrant represent all the basic Pythagorean triples shown in Table 1.

The scatter plots Figure 3 and Figure 4 show that basic Pythagorean triples are very sparse and regularly distributed in comparison with the scatter plot in Figure 1, which means that the latter contains more Pythagorean triples. The missing ones are multiple Pythagorean triples.

4. Scatter plot of all Pythagorean triples

a) Multiple Pythagorean triples

Because horizontal and vertical parabolas give the same Pythagorean triples, we can get all Pythagorean triples of the first quadrant with vertical parabolas only. So, we compute the Pythagorean triples with vertical parabolas and plot the scatter plot in the first quadrant.

A vertical parabola is determined by the value of its parameter d , see (10). For plotting all the vertical parabolas, we will plot a series of parabolas with d ranging from a minimal value d_{\min} to a maximal value d_{\max} , with d_{\min} and d_{\max} defining the smallest and the largest parabola respectively. We set a smallest parabola because near the origin the dots are so close to one another that they are undistinguishable, see Figure 1. So, the parabolas with $d < d_{\min}$ will not be plotted.

- Maximal k

A characteristic of a vertical parabola is the value of Y at the vertex which is given by equations (10) and (21) with $X=0$. This value is the Y_v below :

$$\begin{aligned} Y_v &= -\frac{0^2 - d^4}{2d^2} & (44) \\ &= \frac{d^2}{2} \end{aligned}$$

The largest parabola is defined by d_{\max} and the Y of its vertex is:

$$Y_{\max} = \frac{d_{\max}^2}{2} \quad (45)$$

For multiple Pythagorean triples, the Y of the vertex of a multiple parabola equals kY_v and is smaller than Y_{\max} . So, all k should be such that:

$$kY_v \leq Y_{\max} \quad (46)$$

Let k_{\max} be the maximal value of k , we have:

$$k_{\max} \frac{d^2}{2} = \frac{d_{\max}^2}{2} \quad (47)$$

$$k_{\max} = \left(\frac{d_{\max}}{d} \right)^2 \quad (48)$$

b) Scatter plot

Each basic Pythagorean triple on the parabola d is defined with the parameter i . Because the interval of i for a parabola that spans over the first and second quadrants is given by equation (43), the interval of i for the first quadrant is half that interval which is:

$$-\frac{d}{2} \leq i \leq 0 \quad (49)$$

Let d_{\min} be the value of d of the first parabola. With the following values of the parameters: $d_{\min}=13$, $d_{\max}=100$, $1 \leq k \leq k_{\max}$ and $-\frac{d}{2} \leq i \leq 0$, we have plotted the scatter plot of Pythagorean triples which represents basic Pythagorean triples and multiple Pythagorean triples in Figure 5.

We have also plotted four horizontal and vertical parabolic curves for comparison with the parabolic patterns in the scatter plot. We see that the parabolic patterns match well with the four parabolic curves.

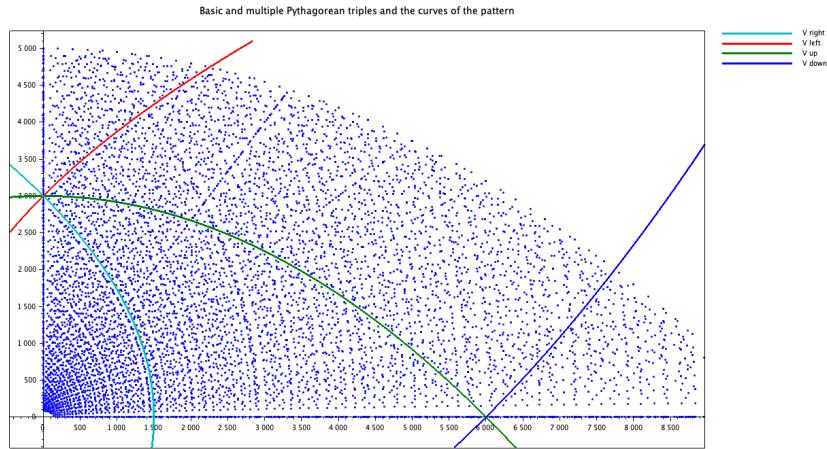


Figure 5

c) Video of the plotting

It is interesting to see the plotting in the making with a video because we see the structures emerge. The video is here: [Plotting of scatter graph of Pythagorean triples](#)

- Description of the video on YouTube

A collection of Pythagorean triples are represented as a cloud of dots on the plane of the coordinate system which is called a scatter plot of Pythagorean triples. Within the scatter graph, there are sets of parabolic patterns with a high density of dots. Here is the plotting process as the Pythagorean triples are being computed and the parabolic patterns emerge.

5. Why the parabolic patterns show up

If a cloud of dots is randomly scattered, there will be no pattern in the cloud. We see parabolas in the scatter plot because the density of dots along the parabolas is higher than its surrounding. There are also white parabolas where the density of dots is lower. We can compute the density of parabolas from the dots on the X-axis because all parabolas intersect with the X-axis. The Y at the intersection points equal 0, which allow us to compute the X for all dots on the X-axis.

Let X_d be the ordinate of the intersection point of the parabola d . With $Y=0$, we compute X_d , see (22):

$$X_d = d^2 \quad (50)$$

Let this parabola be a basic parabola. The ordinate of the intersection point of a multiple parabola is kX_d . So, general parabolas intersect with the X-axis at:

$$X_0 = kX_d = kd^2 \quad (51)$$

with $d=1, 2, 3, 4 \dots, k=1, 2, 3, 4 \dots$.

So, the X_0 for all k and d equals :

$$X_0(k, d) = (1, 2, 3, 4 \dots)(1, 2, 3, 4 \dots)^2 \quad (52)$$

Obviously, these $X_0(k, d)$ are not evenly scattered on the X-axis. The density of dots per unit length at the i^{th} dot equals the number of consecutive dots n divided by the distance between the first and the last dots:

$$D(i) = \frac{n}{X_0(i) - X_0(i - n + 1)} \quad (53)$$

Because one point corresponds one parabola, the density of dots $D(i)$ equals the density of parabolas. We have plotted $D(i)$ in Figure 6 where several high peaks appear. These peaks indicate high density regions which correspond the parabolic pattern in the scatter plot.

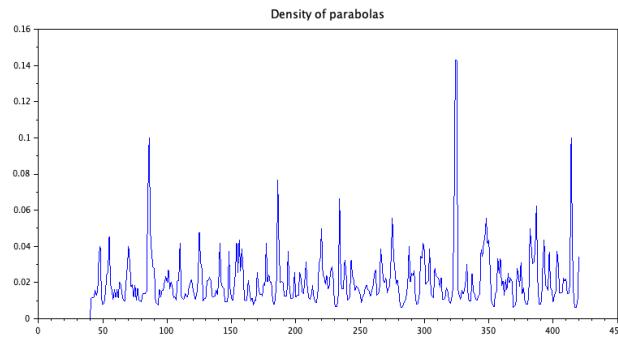


Figure 6

6. Discussion

In this study, four parabolic functions were derived directly from the Pythagorean equation. These functions provide a mathematical description of the parabolic structures observed in the scatter plot of Pythagorean triples.

The scatter plot of basic Pythagorean triples reveals a highly regular arrangement on the (X,Y) plane. The plot of the basic Pythagorean triples shown in Table 1 indicates that the basic triples occur precisely at the intersections of horizontal and vertical parabolas.

Figure 5 illustrates the superposition of the four derived parabolic functions on the scatter plot of all Pythagorean triples. The strong correspondence between the curves and the observed structures confirms the validity of the functional model.

In addition, the density distribution of the parabolas was computed, identifying regions of higher concentration. This analysis clarifies why the parabolic structures are so prominent in the scatter plot of all Pythagorean triples.

Letter to readers

I would like to explain how I came to analyze the parabolic patterns in the scatter plot of Pythagorean triples. After writing «[Classification of Pythagorean triples](#) and [reflection on Fermat's last theorem](#)», I revisited the Wikipedia page on [Pythagorean triple](#) and noticed these parabolic patterns, along with their description. I had not been aware of this feature previously.

The classification of Pythagorean triples provided me with a deeper understanding of the Pythagorean equation and its solutions, which enabled me to recognize these parabolic patterns. Since no explanation for these patterns existed, I realized this was a problem worth addressing.

I recalled the parabolic curves of the classified Pythagorean triples that I had plotted in Figure 1 in my article. The patterns in the scatter plot are also parabolic. A pattern reflects a geometric order of the plotted points, which in this case represent Pythagorean triples I had just classified. This gave me confidence that I could explain these patterns mathematically.

Applying the change of variables to the Pythagorean equation, I immediately obtained four parabolic functions. These functions fit the observed parabolic patterns precisely. Consequently, with the theoretical framework from the classification, explaining the parabolic patterns was straightforward. I would not have achieved these results without this underlying theory.

This experience demonstrates that when a mathematical theory is available, solving related problems becomes much more feasible. Often, solving a problem requires creating the theory that underlies it. This confirms my previous observation that discoveries often arise from the conjunction of multiple unpredictable factors. A possible way to make discovery could be pursuing a challenging objective, which inevitably produces intermediate theories, some of which may appear unrelated but can later prove key to other problems—just as the classification of Pythagorean triples enabled the explanation of the parabolic patterns in the scatter plot.

I do not intend to submit this article for journal publication, not out of reluctance to share it, but because I expect it will be rejected. For example, my article «[Extending complex number](#) to spaces with 3, 4 or [any number of dimensions](#)» was rejected by a mathematical journal with the note:

“Editorial Board does not consider your submission to lie within the scope and standards we currently must set for [name of the journal]. Accordingly, it has been rejected without review.”

I understand this reaction, as I am not a professional mathematician and cannot write in the formal style expected by journals. Nonetheless, I believe my ideas and theories are valid. Examples include the [Classification of Pythagorean triples](#), [Analyze of Cantor's theorem](#) and [uncountability, N-complex number, N-dimensional polar coordinate](#). Repeated rejection is frustrating, and the thought of these advanced theories being lost is disheartening. I would therefore like to request that readers share my work with other mathematicians.

Additionally, I would like to propose a suggestion regarding the naming of theorems or theories by contributors with Chinese names. Chinese family names are typically short, often only one syllable, and common across millions of people. For example, the [Chen's theorem](#) was proven by Chen Jingrun, but the name Chen could be confused with the name of Chen Xingshen (Shiing-Shen Chern in western spelling) who is another famous mathematician. Indeed, the family name Chen is shared by over 54 million people. So, the same family name can lead to confusion regarding attribution. As Chinese family name is short, we add the character 氏 after the family name Chen which makes [Chen's theorem](#) to signify literally theorem of the family Chen. So, I believe a clearer method is to include the full name of the contributor.

For example, a theorem by Chen Jingrun should be named ChenJingrun's Theorem, and a theory by Chen Xingshen should be named ChenXingshen's Theory. This approach provides clear credit to the actual contributor and should become the standard for attributing work by Chinese scholars. In this spirit, if my own work were to be named, I would prefer PengKuan's Theorem or PengKuan's Theory, rather than simply Peng's Theorem.